

## **The Economics of Shame in Work Groups: How Mutual Monitoring Can Decrease Cooperation in Teams**

Shepley W. Orr\*

### **I. INTRODUCTION AND STATEMENT OF THE PROBLEM: THE NEED FOR MONITORING IN AGENCY THEORY**

The principal-agent problem is one of the foremost difficulties in the field of incentive design (Campbell 1995 presents a recent discussion). The problem is one of asymmetric information: because a principal lacks information about the agent's ability or his intention to work, the principal must find some means of discovering or eliciting that information to ensure performance.

Principal-agent theory recognizes many mechanisms for promoting performance, and I will focus on only one here: the use of group piece-rate incentive schemes. Under such schemes, instead of paying an agent for each piece produced, agents are put into groups and paid based on the aggregate output of the group (Petersen 1992a, and see equation 1 in the Appendix)<sup>1</sup>. Because each agent relies on the others for pay, the firm shifts the monitoring burden onto the agents in the group. Before addressing the potential efficacy of monitoring, note that the incentive structure of a group piece-rate scheme is that of a public good, and hence is susceptible to free-riding. Principals must believe, therefore, that mutual monitoring in the group will alleviate the free-rider problem, and ensure performance, in a way that will be more efficient than the alternatives of (1) monitoring by the firm, (2) individual piece-rate incentives, (3) simply using salary pay, or some combination of the three.

Group piece-rate incentive schemes have often been studied by sociologists (Burawoy 1979, Homans 1951, Roy 1953, and Whyte 1955) who generally ar-

\* Department of Sociology and Balliol College, Oxford. I am grateful to Michael Bacharach, Natalie Gold, Gerry Mackie and Avner Offer for helpful comments. Balliol College, Oxford, UK. OX1 3BJ. E-Mail: shepley.orr@balliol.ox.ac.uk.

1. Of course, group piece-rate incentive schemes are structurally similar to gainsharing or profit sharing schemes in general.

gue that the threat of social sanctions from other group members will ensure that each agent contributes his share to the labour input. This assumption, recently renewed in economic models (discussed below), is roughly as follows: being in a group promotes group identity and a sense of having a mutual fate; group identity and mutual fate create the threat and fear of disapproval from other group members; therefore, the task for which the group was created will necessarily be successful, for no single agent will want to deviate from what benefits the group<sup>2</sup>.

It has been argued, in contrast to the sociological assumptions, that not only will the good not be produced for game-theoretic dominance reasons which predict no contributions, but nor will anyone undertake sanctioning, for social sanctioning is merely another public good of a second-order, and the Nash equilibrium is to neither contribute to the public good nor to sanctioning others (see Boyd and Richerson 1992, Calvert 1995, Heckathorn 1993, and Sethi and Somanathan 1996 for formal models of sanctioning<sup>3</sup> in public good problems and discussions of conditions under which sanctions in public goods problems might emerge). I too will argue that mutual monitoring is potentially ineffective, but not from the view that egoists will fail to cooperate on the first- or second-order public goods problem. Nor will I follow Hechter (1987) in assuming that all groups require formal organization to undertake the monitoring and sanctioning needed for ensuring cooperation. In contrast to standard economic assumptions, I will employ psychologists' model of *hidden costs*, or what is known in economics as the *motivational crowding effect* (Frey 1997a, Frey and Jegen 2000) to explain the potential failure of mutual monitoring. On the assumptions of motivation crowding theory, agents can be intrinsically, and specifically in our case *normatively*, motivated to 'do their share' in contributing to a public good, but excessive external intervention which is perceived as controlling may lead to a decline in their motivation to act cooperatively<sup>4</sup>. Much work on crowding has dealt with relationships between principals and agents, but none has shown how principal's behaviour can lead to motivational crowding within a group. In this essay I will discuss how the implicit presence of the price mechanism in social relations in groups may lead to a similar crowding-out of the normative motivations which lead to cooperation in teams.

2. See Hechter (1987) for a critique of this position in sociology which he calls 'normativist', and which more generally is known rhetorically as 'homo sociologicus'.
3. I use 'sanction' here as synonymous as with 'punishment'.
4. The position here then differs from an unpublished paper by Bowles and Gintis (1998) in which the motivation to monitor comes from the fact that benefits are shared and that agents have a 'social preference' (1998, p. 4; i.e., a non-individual-wealth-maximizing preference) to reciprocate the efforts of others. I argue that the fact of shared interests where individuals act on such

## II. SOCIAL REWARDS AND SANCTIONS IN WORK GROUPS: PREVIOUS ECONOMIC MODELS

Recently, economists have pointed out that not just pecuniary and disciplinary punishments can serve to reinforce desired patterns of action, but also that social sanctions and social disapproval can be shown to have an effect on behaviour. I will make two critiques in this Section against these models. One critique addresses the idea, recently put forward by some economists, that egoistically motivated agents can effectively socially sanction one another, which I argue is theoretically incoherent. The second critique is that economic models that accept that social sanctions *can* effectively bring about desired conduct have tended to mis-specify the nature of the utility for approval that leads actors to respond to social sanctioning.

### 1. *An A Priori Argument against Theoretical Derivations of Social Sanctions from Egoism*

Before addressing the particular economic models that in different ways attempt to show how egoistic agents sanction one another and accept social sanctions, I shall make a very brief *a priori* argument against this very idea. Take the following example: suppose that you and I are two egoists, and have made no agreements to cooperate in a prisoner's dilemma-type situation. As the standard economic model predicts, neither of us cooperate. Now suppose that you decide that it will be in your interest to disapprove of, or socially sanction, me to get me to cooperate. It is common knowledge that we are both egoists. Your problem in attempting to socially sanction me is to invent any kind of content to your sanction which is not transparently self-effacing. For, what could you possibly say to me that would switch my behaviour into the desired direction if you are an egoist and I am as well? Your emotional reaction cannot be

social preferences makes mutual monitoring unnecessary, and hence can lead to crowding out on the part of those who are motivated to act on social preferences. Also, Bowles and Gintis (1998) do note a distinction between altruism and reciprocity and mention Rotemberg's model (discussed below in II.2), and note that 'were team members sufficiently altruistic in this sense to motivate mutual monitoring, there would be no initial free-rider problem either' (p. 4). However, the distinction between altruism and a social preference is not clarified. Presumably, a social preference is a preference for an inequality minimizing distribution, whereas altruism solely concerns the degree to which other agents are benefited. If so, then I agree with their critique of Rotemberg, but their argument about the superfluosity of monitoring under altruistic preferences would also apply to their suggested case of reciprocity motivations based on a social preference.

anything other than anger that I did not happen to act so as to put more money in *your* pocket (Rawls 1971, p. 488). Even if I were non-egoistic and could accept an appeal to some value that might make me feel guilty or ashamed for non-cooperation, nothing that you as an egoist can say will motivate me to change my behaviour, as you have not pointed out any fault in my character or a failure to act according to a standard that you yourself accept (or can understand), as you have no such standards, nor any values which uphold them<sup>5</sup>. Conversely, if I am non-egoistic and you are egoistic, there is nothing I can say to you that will persuade you that you should comply with an obligation. The evaluative content needed to deliver or accept a social sanction that is even intelligible, let alone persuasive, is unavailable to any egoist, and therefore any economic model which attempts to show that rational egoists can create cooperative behaviour through *social* sanctioning is necessarily mistaken<sup>6</sup>.

## *2. Specific Arguments against Theoretical Derivations of Social Sanctions from Egoism*

I address three such proposals for the evolution of effective social sanctions from egoism here. First, it has been suggested that social sanctions would be effective because agents have an intrinsic utility to conform to social norms (Jones 1984). In addition to the fact that this model effectively dissolves the problem it is meant to address<sup>7</sup>, we may note Sugden's (1985) critique of this argument that an intrinsic utility to *conform* is quite different from an intrinsic utility to respond to *social pressures to conform*. If all agents had intrinsic utility to conform there would be no need for social pressures at all; at worst hu-

5. Of course, it may be that the egoist's criticism leads to me feeling guilty simply because I know that I may have acted so as to violate a standard I hold: but this is not the egoist *making me* feel guilty; he merely reminds me of something that *I make myself* feel guilty about. That is, it is *not* the egoist's *disapproval* that leads to my guilt, for egoists by definition cannot express moral disapproval.
6. I do not defend this claim further here, but refer the reader to chapter one of my doctoral thesis (Orr 2000), and to arguments made by political philosophers which stress three points: (1) approval cannot be a direct function of price (Pettit 1990); (2) disapproval among egoists is not possible because approval and disapproval are moral attitudes which presume that agents are motivated by concerns of fairness and justice, which by definition are precluded for egoistically motivated agents (Rawls 1971, p. 488); (3) social sanctioning as moral condemnation, therefore, is *unintelligible* in the context of the egoistic wealth-maximizing assumption (Barry 1995, pp. 35–36).
7. And is subject to the well-known 'functionalist fallacy': if humans want to conform, then the existence of norms is a functional consequence of human nature. Such 'explanations' fail to explain.

mans would face coordination problems concerning which norms to adopt. If we have intrinsic utilities to respond to social pressure to conform, we must conform because we are given reasons for why some norm is good for the agent, for the collective, or some other normative reason, which necessarily will be exogenous to the model (Kuran 1995).

A second model is that in which ‘peer pressure’ is explained by the fact that

‘as long as a worker is told that he is to punish the neighbor on his right or suffer punishment from the one on his left, he will carry out the punishment’ (Kandel and Lazear 1992, p. 813)<sup>8</sup>.

Of course, the problem here is that this merely restates the problem of collective action but at the ‘second-order’ level (i.e., why not free ride on other people punishing non-cooperators, and why not free ride on the efforts of those who punish those who fail to punish those who don’t cooperate, and so on?). Some formal solutions to the punishment problem have been proposed in which costs of sending and receiving punishment and levels of interest and information vary (Boyd and Richerson 1992, Calvert 1995, and Heckathorn 1993, Sethi and Somanathan 1996), but all such models must either (i) assume that some agents are already normatively motivated, in which case they are subject to the problems discussed in III.2 below, or (ii) are subject to the *a priori* criticism of egoism and social sanctions that I have made above.

A similar standard economic approach suggests that a prosocial motive can be acquired in the form of an *altruistic* interest in others from selfish motives (Rotemberg 1994). Here, the worker who does not have any sentiment for other workers, and hence knows that he is unlikely to have that other worker reciprocate work effort, can effectively force himself to like the other agent, because he realizes that this will be economically advantageous for himself. Rotemberg suggests that we can acquire a ‘taste’ for altruism merely by spending time with other agents, much in the same way that an agent can acquire a taste for bourbon if one drinks it often enough. Two problems arise here. First, it assumes that an agent can come to have positive regard for anyone, and this is an implausible

8. More specifically, Kandel and Lazear state: ‘If there is an end to the chain so that someone believes that he will not be punished for not disciplining another who let another go without punishment, then there is no reason for anyone to discipline anyone’ and then derive their model of mutual punishment by simply postulating ‘But the firm can be thought of as a circle. As long as a worker is told only that he is to punish the neighbor on his right or suffer punishment from the one on his left, he will carry out the punishment’ (1992, p. 813). The ‘circle’ model of the firm though, itself stands in need of explanation in order to know what supports the beliefs about whether others will punish or not, and *how* the sanctioning crowds-in a normative motivation. See Barron and Gjerde (1997) for a further modification of Kandel and Lazear (1992) which models principal’s behaviour.

assumption<sup>9</sup>. Second, this model assumes that the agent can actually bring himself to adequately fake his esteem for another agent. Although falsified approval is surely possible and perhaps occurs often, it should not be taken as the ideal type. As a definition of friendship it is mistaken (Lane 1991), and further such ‘pseudo-regard’ is likely to be detected (Offer 1997).

In sum, although these models may be more parsimonious as they rely on the standard economic assumptions, their conceptual implausibility is an argument for slightly more complex, but more realistic, models of social sanctioning in work groups.

### III. SOCIAL RELATIONSHIPS AS SUPPORTIVE OR CONTROLLING EXTERNAL INTERVENTION

Behind these criticisms lie a few simple assumptions about how social relationships motivate human agents to act on their normative beliefs. First, humans cannot approve of one another based on price alone. Second, the mere fact of another agent’s disapproval is a necessary but insufficient factor in explaining a change in behaviour. Here I will focus on how mutual monitoring may itself be a form of disapproval which affects one’s normative motivations. This requires a brief discussion of motivation crowding theory (Section III.1), and in Section III.2 I will show how the assumptions of motivation crowding theory suggest that mutual monitoring can decrease normative motivation. Section III.3 discusses different responses to monitoring based on motivational type.

#### *1. Motivation Crowding: Controlling vs. Supportive External Interventions*

I will here provide only a very brief summary of motivation crowding theory, and refer the reader to other sources for a more thorough discussion and bibliography (Frey 1997b, Frey and Jegen 2000).

When principals wish to bring about some form of behaviour in agents, it is assumed that some form of action by the principal must be taken beyond merely contracting for the agent to take the required action. The principal’s actions can be in the form of rewards or punishments, which may come in a variety of forms but economists focus on the use of pecuniary incentives. Agents may or may not be intrinsically motivated to undertake the task, and they may

9. Rotemberg restates the assumption of Homans in *The Human Group* (1951) that sentiment for others can be generated merely by participating in shared activities with others. Simple introspection into experience shows this to be an unhelpful assumption: it suggests that individuals, on average, like people with whom they spend time simply because they spend time with them.

be intrinsically motivated by virtue of either (i) enjoying the task itself or (ii) because they are normatively motivated to comply with their obligations.

Rational choice theory assumes that if an agent is willing to undertake some action initially, then offering an incentive to undertake the task or a punishment for failure to undertake the task can only bring about more of the required action. Motivation crowding theory, in contrast, suggests that such external intervention may lead to an ‘overjustification’ of the action for the agent, and results in the agent feeling either (i) a diminished sense of self-esteem concerning his ability (where the principal lacks confidence in the agent’s skills), or (ii) that he is not trusted to take the action without monitoring or incentives.

As an example of monitoring crowding-out intrinsic task (and/or perhaps normative) motivation, consider Barkema’s (1995) result that managers who are more extensively monitored in fact perform less than those in whom greater trust is shown (see Frey 1993 for a theoretical model). As an example of crowding out normative motivation by use of the price mechanism, take Gneezy and Rustichini’s (forthcoming) study of the use of fines for parents who picked up their children late from day-care centres. In the group that was fined, the lateness of picking up children increased, and stayed at the same level when the fining was removed; a control group with no fining stayed at the same level throughout the same period.

Not all forms of external intervention lead to crowding-out, however. When intervention is seen as either *supportive* or as recognizing the agent’s *general capacities* (see Frey 1997b), then the intervention leads to crowding-in intrinsic motivation.

In short, when a principal provides excessive rewards or price incentives or other forms of disciplinary intervention (e.g., monitoring, imprisonment, fining) the agent either changes his preferences or changes the way he perceives the situation (Frey and Jegen 2000, pp. 6–7). I am concerned with the situation in which a change in frame (i.e., perceiving the situation) occurs by virtue of monitoring, as described by Lindenberg and Frey (1993):

‘Frame switches have . . . been observed . . . where monetary rewards displace a normative frame by a gain frame’ (Lindenberg and Frey 1993, p. 200)<sup>10</sup>.

10. Pettit (1996) argues similarly for the frame switching effect of excessive disciplinary intervention, but not for reasons of diminishing the agent’s self-esteem. Instead, Pettit argues that the existence of excessive discipline and monitoring will signal distrust in the population, leading agents to believe that it must be the case that most other agents will not comply. Not wanting to be ‘suckers’, agents who would otherwise cooperate then choose not to. As Pettit puts it: ‘the introduction of deviant-centered sanctions would tend to switch agents from a nonegocentric to an egocentric mode of deliberation’ (1996, p. 73; for a similar argument see Frey 1997a, and Bowles 1998, p. 105).

My argument is that what I will call the suspicion effect below is a result of a change in frame due to the regime of mutual monitoring, in which agents are likely to switch their framings of situations from normative to gain frames; this leads agents to see what might be normal acts of friendship as instead acts of monitoring.

## 2. *Mutual Monitoring as Distrusting and Controlling Intervention*

To make the crowding out model clear with regard to mutual monitoring, two assumptions must be stated at the outset. First, agents must be assumed to be capable of some kind of moral argument or reasoning about obligations which is the content of rational social persuasion. Second, it must be assumed that agents are normatively motivated to differing degrees.

It is once we assume *varying* degrees of intrinsic normative motivation in a population of agents that the problem of mutual monitoring becomes interesting, and problematic<sup>11</sup>. The problem has two components, which result from the presence of the price mechanism in the formation of the group. First, because principals effectively assume that agents will become friends (i.e., have positive affect for one another, like one another, or what have you) such that social sanctioning can ensure performance, then the friendships which may ensue may seem to have a coerced element to them. Call this the *manipulation effect*. Second, as a consequence of the manipulation effect, if friendships do in fact emerge, it may be the case that agents will necessarily have the worry that the friendships are, if only in part, motivated by egoistic motives, in which the egoistic agent fakes his friendship so that he will achieve optimal pay. This is precisely the coerced friendship that Rotemberg suggests the agent can force himself to adopt, and thus fails to note that friendships lose value when friendships are externally coerced or are undertaken for non-intrinsic reasons (Lane 1991). Call this the *suspicion effect*.

A study by Enzle and Anderson (1993) illustrates in the laboratory how the suspicion effect may decrease intrinsic motivation in work groups. They had subjects perform tasks under no surveillance, and in situations where subjects believed that monitoring was non-controlling or controlling. The intentions attributed to monitoring impacted upon performance: it was found that ‘controlling’

11. Bohnet, Frey and Huck (2000) assume different types and model effects of varying probabilities of contract enforcement and varying degrees of severity of enforcement in the form of fines. They do not control, however, for the initial conditions, viz. the *proportion* of cooperative to individualistic types (what they call Honest and Money-maximizing types), which will accelerate the rate of motivational crowding, so I argue.



surveillance yielded the lowest intrinsic motivation. Further, when subjects did not know why they were being monitored, they were likely to think that distrust motivated the monitoring. Similarly, in real world work groups, the suspicion effect suggests that agents will be more likely to attribute controlling surveillant intentions to other agents, and this will lower their own intrinsic motivation.

The suspicion effect is most harmful when the suspicion is well-founded; that is, when some agents are egoistic and are likely to use friendships to ensure performance in their co-workers. An agent is likely to respond to intrusive monitoring as did one of Whyte's (1955) respondents in a work group under group piece-rate incentives:

'Why the hell should I care what a bunch of damn snoopy bums think of me?' (cited in Levine 1992, p. 482).

It is possible that *because* the agent thinks his co-workers are 'snoopy', he will then not care what they think of him. An environment in which 'snoopiness' is encouraged by the price mechanism may then encourage agents to not care what other agents think, and decrease the value of social rewards.

Let me illustrate the suspicion effect with an example. Suppose that A is normatively motivated, and A believes that perhaps B is not. When A is working and B is behaving in a friendly manner to him, A becomes suspicious of B's motives because A feels that B's behaviour is in the service of monitoring, and not the behaviour of a friend<sup>12</sup>. If so, then A will therefore see the normative motivation to cooperate as being *coerced* by the monitoring of B, and change his framing of the choice to cooperate from being a normative obligation to being a market relationship. Further, if A resents B's 'snoopiness', A will reduce his output partly because he has no desire to benefit B.

The problem is that the suspicion effect can lead to false attributions of monitoring intentions to other agents: the implicit presence of the price mechanism in social relations as a consequence of the mutual monitoring regime may make agents more likely to attribute (egoistic) monitoring intentions to otherwise benign behaviour. These patterns of attribution have a negative effect by both (a) increasing suspicion that the agent himself is not trusted by fellow agents which crowds out his normative motivation to comply, and (b) by reducing his desire

12. Assume that observability of behaviour is not chosen, as is likely to be the case where agents are working in closed quarters, or where task interdependence ensures that agents view each others work. What is chosen by the agent is whether or not to monitor, which is not necessarily behaviourally distinct from other activities; thus, I suggest agents will attribute monitoring *intentions* or not as a function of the 'attributing' agent's type and whether or not mutual monitoring is encouraged.

for the positive externalities of his actions to benefit his fellow agents (e.g., by increasing the pay to each agent).

### 3. Differentiating Responses to Monitoring Based on Motivational Types

Let me summarize the argument made here by offering some cognitive mechanisms which may explain how normative motivation is crowded-in or – out by monitoring. Let us assume simply that there are two types: normatively motivated or ‘cooperative’ types, and egoistic or ‘individualistic’ types.

To begin, question whether the effect of monitoring has a *strictly linear* crowding-out effect in normatively motivated types. I believe not. While an agent may be intrinsically normatively motivated, this does not entail *aversion* to being noticed or praised for taking normatively motivated actions. Motivation crowding-theory suggests that there are diminishing and eventually negative returns to being noticed or praised: too much praise for something one would do anyway is patronizing; too much of being noticed for something one would do anyway signals a lack of trust, and this ‘noticing’ of actions *becomes* monitoring. It is for this reason I claim that the effect of mutual monitoring on the degree of normative motivation in cooperative types is non-linear: while the monitoring is initially crowding-in of normative motivation, as being noticed acting on one’s ‘better’ motives raises the agent’s pride in his action, at some point the normatively motivated agent begins to perceive the monitoring as distrustful or coercive, and his normative motivation decreases. This suggests that the maximum in the curve representing effort responses to monitoring is the point where one frames what was ‘noticing’ as now ‘intrusive monitoring’, and produces a corresponding shift between a normative and a gain frame (see *Figure 1* below, and the quote by Lindenberg and Frey 1993 above).

A second question is whether monitoring actually crowds-*in* motivation in mostly egoistic types? I believe the answer may be yes, as surely the non-normatively motivated are occasionally made to develop normative motivations. I will offer three possible cognitive mechanisms as to why crowding-in may occur in individualistic agents.

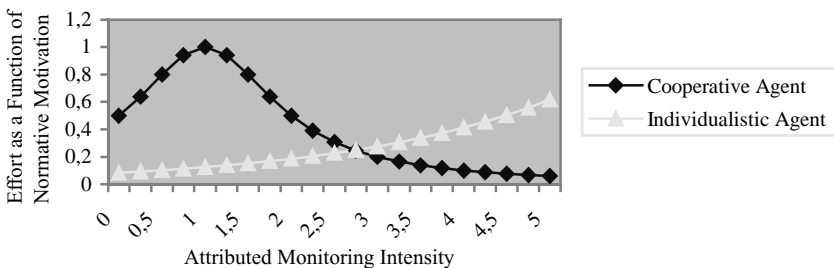
One is that agents with relatively individualistic/selfish orientations require monitoring to realize that the obligation to comply with the norm is serious. Individualists tend to believe that other people are all selfish like themselves (see Kelley and Stahelski 1970). When an individualist sees others complying and sanctioning others for not complying, the individualist may come to see the other agents not as being ‘suckers’, but instead may revise her beliefs about the possibility of agents being normatively motivated. This is a form of normative learn-

ing in which the frame shifts from a ‘gain’ to a ‘normative’ frame, by seeing that other people are committed to undertake the effort of monitoring and sanctioning. This is particularly true for individualists who were non-cooperative for *fear* of being a sucker (and not out of *greed*; see Yamagishi and Sato 1986).

A second cognitive mechanism supporting the view that monitoring may have a positive crowding-in effect for individualistic types is that such agents may devise self-justifications which allow them to frame their actions as intrinsically versus extrinsically motivated. Because the agent is being monitored such that shirking is no longer optional, the agent must devise an intrinsic attribution instead of the extrinsic intervention of fear of punishment to justify to himself why he is working (cooperating). Hence, the agent self-servingly attributes his action to a normative motivation, and in so doing, must think about normative motivations, which is likely to lead the agent to take such normative reasons for acting seriously (for a similar crowding-in explanation see Akerlof and Dickens 1982 and Dickens 1986 on how lessened punishments encourage moral reasoning about why to obey the law).

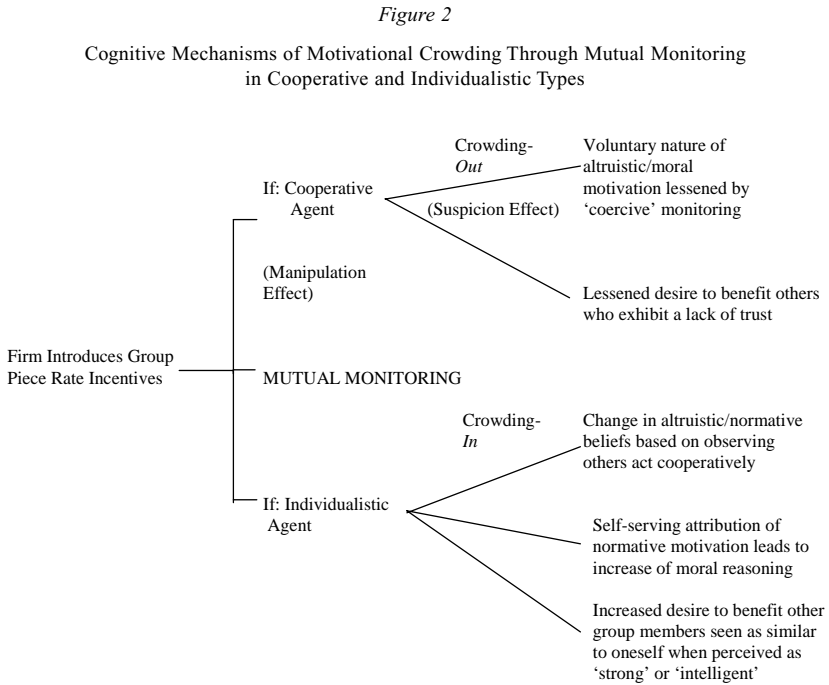
A third cognitive mechanism for why social sanctions and monitoring can increase normative motivations in individualistic types is that egoistic agents frame non-compliance as ‘intelligent’, and are more likely to respect one another as being ‘shrewd’ or ‘strong’ (Liebrand et al. 1986, Van Lange and Liebrand 1991). If we assume that agents do want to benefit others that they perceive as like themselves, and group incentive schemes deliver benefits to the other agents, then monitoring provides evidence that other agents are similar, and cooperating will increase the amount of utility derived from benefiting those in the group.

Figure 1  
Crowding Out Normative Motivation



If these arguments are correct, then we can graph the crowding-out of normative motivation in cooperative (normatively motivated) types and the crowding-in of normative motivation in individualistic types as in *Figure 1*. As suggested, for cooperative types, as monitoring increases there exists a maximum at which the frame shifts from a normative to a gain frame. For individualists, monitoring will tend to crowd-in normative motivation in a more linear manner<sup>13</sup>.

The argument can be summarized in terms of a set of cognitive mechanisms as in *Figure 2* below:



13. In the Appendix below, I defend the mathematical assumptions in *Figure 1* by postulating a utility function that responds to overall levels of monitoring and/or disapproval.

#### IV. CONCLUSION: IMPLICATION FOR RESEARCH ON INCENTIVES IN WORK GROUPS

I have shown here briefly that the optimism shown by economists for the role that social sanctions can play in situations of mutual monitoring is misplaced, and that social sanctions require that agents have some intrinsic normative motivation. Hence, social sanctions cannot be derived endogenously from assumptions of self-interested motivation. Petersen (1992b) has suggested that *both* normative (moral) motivation and social rewards from cooperation and fear of disapproval for non-cooperation are likely to lead to greater cooperation under group piece-rate incentives. This is an important step in modifying the standard economic assumptions of behaviour in firms, but should be supplemented by the analysis presented here, in which social rewards and sanctions can interact in a non-linear negative manner with extant normative motivations to produce less cooperative behaviour.

What implications for future research follow from the arguments made here? First, it must be accepted that there are no policies of monitoring or sanctioning in groups that will be effective independent of the types in the population. If there exists excessive reliance on monitoring and discipline where there are mostly normatively motivated types, the effects may be negative for performance. Conversely, if there exists no monitoring and discipline where agents are mostly individualistic, the result may be complete free-riding. Motivational crowding is, in the language of evolutionary game theory, frequency dependent.

Further, because mutual monitoring among cooperative types decreases the value of their normative motivations through the manipulation and suspicion effects, two counter-intuitive implications can be derived. Our second implication for research, then, is as follows: in contrast to the usual assumption that there is a negative relationship between group size and levels of contributions to public goods, larger groups may actually be more productive if the group contains many normative types. Larger groups will allow individuals more freedom to choose with whom they associate, and decrease the manipulation effect of feeling that friendships are coerced by the firm. Thus, a larger group size crowds-in the intrinsic nature of friendship-selection.

A third implication is that where (i) piece rate incentives and not salaries are preferred, and (ii) agents are mostly normatively motivated, it may be best to abandon *group* incentive schemes altogether. If the point of group piece-rate incentives is to enable inexpensive monitoring for the firm, and this mutual monitoring leads to adverse consequences in cooperative types, then it may be more productive for the firm to simply employ individual piece-rate incentives.

Firms may do best to simply rely on the type of ‘gift-exchange’ (Akerlof 1982) and/or trust incentives between principal and agent to do the work of ensuring performance.

The upshot of these implications is that they necessitate the identification of individual types. Further social scientific theory and research is required in order to advance both the understanding of motivational types and their interaction with one another under different policies of compensation, monitoring, and discipline.

## APPENDIX

### *Modifying Hollander’s Social Exchange Equilibrium to Account for Motivational Crowding*

Hollander’s (1990) model<sup>14</sup> of a ‘social exchange’ explanation of voluntary contribution to public goods states that an agent prompts from the other group members a ‘sentiment bundle’ of emotional reactions, which are positive (approval) or negative (disapproval), when he voluntarily contributes to a public good. This is a function of  $s$  on  $b$  ( $f = s(b)$ ), where  $b$  is the amount contributed by the agent to the public good. On the usual assumption that there is a production function which determines the marginal per capita return on a contribution to a public good (some multiplicative constant  $\beta$ ), and assuming each agent has some endowment  $e$ , then an agent  $i$ ’s share of the total amount of the public good produced is:

$$\pi_i = e_i - b_i + ((\sum b_{i,j}) \beta) / n \quad (1)$$

Where  $e$  is the initial endowment,  $j$  stands for all other agents in the group, and  $n$  is the total group size. The agent always prefers more approval to less (i.e., for any  $s^l$  and  $s^o$ ,  $s^l$  is preferred if  $s^l \geq s^o$ ). Hollander also assumes that the value of approval received for contributing is a function of not simply the absolute level of approval, but also of comparative approval. Therefore, in addition to the absolute amount of approval  $s(b)$ , one has a weighted comparative approval as

14. I alter Hollander’s formulae slightly for purposes of clarity and consistency. I refer the reader to his original paper for optimization and equilibrium conditions. Hollander models a motivational crowding effect for when the state promotes collective goods, but does not show the differential effects of motivational types on crowding.

a function of whether  $s(b) - s(c) \geq 0$ , where  $c$  is the average contribution. Here Hollander constructs a weighted overall approval value,  $a$ , which is determined by some value  $\alpha$  that reflects how much weight an agent puts upon receiving a high level of comparative approval:

$$a = (1-\alpha) s(b) + \alpha (s(b) - s(c)), \quad 1 \geq \alpha \geq 0 \quad (2)$$

or,

$$a = s(b) - \alpha s(c) \quad (3)$$

One therefore has a utility function in which agents maximizes the expected utility of (1) their own private consumption of the public good  $\pi$ , and (2) the overall weighted value of approval,  $a$ , thus:

$$u_i = u_i(\pi) + u_i(a) \quad (4)$$

My modification is a rather simple one, which distinguishes individualistic and cooperative types on the basis of their response to the aggregate level of approval (i.e., the variable  $c$ ). Recall that in the case of disapproval, if there is too much potential disapproval and mutual mistrust present, the value of that disapproval is lowered for cooperative types, and increased for individualistic types. I represent this differential response to attributed levels of mutual monitoring (and social disapproval) by introducing a parameter  $\lambda$  that varies as a function of the agent's motivational type. Therefore,  $\lambda$  will be an increasing multiplicative constant on  $a$  as a function of  $c$ . That is, the overall utility of absolute and comparative approval (i.e.,  $a$ ) is a function of the overall amount of disapproval occurring in the group (i.e.,  $c$ ). Thus, we modify the utility function to be:

$$u_i = u_i(\pi) + u_i(\lambda(a)) \quad (5)$$

For cooperative types, who initially respond positively to mutual monitoring and expressions of disapproval for low contributions, but then respond negatively, crowding out occurs for that agent when  $c$  reaches some level which I will call  $\mu$ . It is when the level of attributed monitoring reaches  $\mu$  that the cooperative type switches his frame from a normative to a gain frame. Assuming that  $\lambda_{t0} = 1$  for all agents, then for cooperative agents,

$$\lambda = (\lambda_{t-1} + 0.1) \quad (6)$$

for each unit increase in the value of  $c$  where  $c \leq \mu$ , and

$$\lambda = (\lambda_{t-1} - 0.1), \text{ if } c > \mu \quad (7)$$

The individualistic agent's response to the overall value of approval (i.e.,  $c$ ) can be simply represented by the increasing convex function that is described in equation (6) above. For individualistic types, as can be seen from figure one, we assume that the intercept that represents their initial contribution level before attributing monitoring intentions to others is lower than that of cooperators, and so in terms of utility, where an agent  $i$  is an individualist and  $j$  is a cooperator, we state:

$$u_j(\pi) \geq u_i(\pi) \quad (8)$$

This condition can be reversed and one may actually have individualistic agents contributing more to the public good if the number of individualistic agents increases past a given level. This is so as the individualist responds consistently positively to social pressure and mutual monitoring by increasing his contribution.

#### REFERENCES

- Akerlof, George A. (1982). Labor Contracts as Partial Gift Exchange, *Quarterly Journal of Economics*. 97: 543–569.
- Akerlof, George A. and William H. Dickens (1982). The Economic Consequences of Cognitive Dissonance, *American Economic Review*. 72: 307–319.
- Barkema, Harry G. (1995). Do Top Managers Work Harder When They Are Monitored?, *Kyklos*. 48: 19–42.
- Barron, John M., and Kathy P. Gjerde (1997). Peer Pressure in an Agency Relationship, *Journal of Labor Economics*. 15: 234–254.
- Barry, Brian (1995). *Justice as Impartiality: A Treatise on Social Justice, Volume II*. Oxford: Oxford University Press.
- Bohnet, Iris, Bruno S. Frey, and Steffen Huck (2000). More Order with Less Law: On Contract Enforcement, Trust, and Crowding. Working paper, Kennedy School of Government, Harvard University.
- Bowles, Samuel (1998). Endogenous Preferences: The Cultural Consequences of Market and other Economic Institutions, *Journal of Economic Literature*. 36: 75–111.
- Bowles, Samuel and Herbert Gintis (1998). Mutual Monitoring in Teams: The Effects of Residual Claimancy and Reciprocity. Santa Fe Institute Working Paper #98–08–074E. ([http://www.unix.-oit.umass.edu/~gintis/mumonit\\_abst.html](http://www.unix.-oit.umass.edu/~gintis/mumonit_abst.html)).
- Boyd, Robert and Peter J. Richerson (1992). Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups, *Ethology and Sociobiology*. 13: 171–195.
- Burawoy, Michael (1979). *Manufacturing Consent: Changes in the Labor Process under Monopoly Capitalism*. Chicago: University of Chicago Press.
- Campbell, Donald E. (1995). *Incentives: Motivation and the Economics of Information*. Cambridge: Cambridge University Press.
- Calvert, Randall (1995). Rational Actors, Equilibrium, and Social Institutions, in: Jack Knight and Itai Sened (eds.), *Explaining Social Institutions*. Ann Arbor, MI: University of Michigan Press: 57–94.



- Dickens, William H. (1986). Crime and Punishment Again: The Economic Approach with a Psychological Twist, *Journal of Public Economics*. 30: 97–107.
- Enzle, Michael E. and Sharon C. Anderson (1993). Surveillant Intentions and Intrinsic Motivation, *Journal of Personality and Social Psychology*. 64: 257–266.
- Frey, Bruno S. (1993). Does Monitoring Increase Work Effort? The Rivalry with Trust and Loyalty, *Economic Inquiry*. 31: 663–670.
- Frey, Bruno S. (1997a). A Constitution for Knaves Crowds Out Civic Virtues, *The Economic Journal*. 107: 1043–1053.
- Frey, Bruno S. (1997b). *Not Just for the Money: An Economic Theory of Personal Motivation*. London: Elgar.
- Frey, Bruno S. and Reto Jegen (2000). *Motivation Crowding Theory: A Survey of Empirical Evidence*. Working Paper No. 49, Institute for Empirical Research in Microeconomics, University of Zurich.
- Gneezy, Uri and Aldo Rustichini (forthcoming). A Fine is a Price, *Journal of Legal Studies*.
- Hechter, Michael (1987). *Principles of Group Solidarity*. Berkeley, CA: University of California Press.
- Heckathorn, Douglas D. (1993). Collective Action and Group Heterogeneity: Voluntary Provision versus Selective Incentives, *American Sociological Review*. 58: 329–350.
- Hollander, Heinz (1990). A Social Exchange Approach to Voluntary Cooperation, *American Economic Review*. 80: 1157–1167.
- Homans, George C. (1951). *The Human Group*. London: Routledge & Keegan Paul Ltd.
- Jones, Stephen R. G. (1984). *The Economics of Conformism*. Oxford: Basil Blackwell.
- Kandel, Eugene and Edward P. Lazear (1992). Peer Pressure and Partnerships, *Journal of Political Economy*. 100: 801–817.
- Kelley, Harold H., and Anthony J. Stahelski (1970). Social Interaction Basis of Cooperators and Competitors' Beliefs About Others, *Journal of Personality and Social Psychology*. 16: 66–91.
- Kuran, Timur (1995). *Private Truth, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.
- Lane, Robert (1991). *The Market Experience*. Cambridge: Cambridge University Press.
- Levine, David I. (1992). Piece Rates, Output Restriction, and Conformism, *Journal of Economic Psychology*. 13: 473–489.
- Liebrand, Wim B. G., Ronald W. Jansen, Victor M. Rijken, and Cor J. Suhre (1986). Might Over Morality: Social Values and the Perception of Other Players in Experimental Games, *Journal of Experimental Social Psychology*. 22: 203–215.
- Lindenberg, Siegwart and Bruno S. Frey (1993). Alternatives, Frames, and Relative Prices: A Broader View of Rational Choice Theory, *Acta Sociologica*. 36: 191–205.
- Offer, Avner (1997). Between the Gift and the Market: The Economy of Regard, *Economic History Review*. 50: 450–476.
- Orr, S. W. (2000). *The Economics of Shame: Effects of Values and Price on Sanction Efficacy in Social Dilemmas*. Unpublished Doctoral Dissertation, University of Oxford.
- Petersen, Trond (1992a). Individual, Collective, and Systems Rationality in Work Groups: Dilemmas and Market-Type Solutions, *American Journal of Sociology*. 98: 469–510.
- Petersen, Trond (1992b). Individual, Collective, and Systems Rationality in Work Groups: Dilemmas and Non-market Solutions, *Rationality and Society*. 4: 332–355.
- Pettit, Philip (1990). *Virtus Normativa: Rational Choice Perspectives*, *Ethics*. 100: 725–755.
- Pettit, Philip (1996). Institutional Design and Rational Choice, in: Robert E. Goodin (ed.), *The Theory of Institutional Design*. Cambridge: Cambridge University Press: 54–89.
- Rawls, John (1971). *A Theory of Justice*. Cambridge, MA: Harvard Belknap Press.
- Rotemberg, Julio J. (1994). Human Relations in the Workplace, *Journal of Political Economy*. 102: 684–717.
- Roy, Donald (1952). Work Satisfaction and Social Reward in Quota Achievement: An Analysis of Piecework Incentive, *American Sociological Review*. 18: 507–514.

- Sethi, Raji and E. Somanathan (1996). The Evolution of Social Norms in Common Property Resource Use, *American Economic Review*. 86: 765–788.
- Sugden, Robert (1985). Review of: *The Economics of Conformism*, S. R. G. Jones, *Economic Journal*. 95: 502–504.
- Van Lange, Paul A. M. and Wim B. G. Liebrand (1991). Social Value Orientation and Intelligence: A Test of the Goal Prescribes Rationality Principle, *European Journal of Social Psychology*. 21: 273–292.
- Whyte, William F. (1955). *Money and Motivation*. New York: Harper and Row.
- Yamagishi, Toshio and Kaori Sato (1986). Motivational Bases of the Public Goods Problem, *Journal of Personality and Social Psychology*. 50: 67–73.

#### SUMMARY

Recent economic theory suggests that free riding under group piece-rate incentive schemes can be alleviated by mutual monitoring and social sanctioning. This article challenges this assumption by showing that the presence of the price mechanism in mutual monitoring and sanctioning can decrease the motivation to cooperate for certain types of agents: because the social rewards for cooperation that may develop through work are potentially based in a desire for pecuniary gain, withholding approval may matter less to initially cooperative agents. Hence, mutual monitoring can decrease cooperation in teams. The author presents evidence from social psychology illuminating differences between individualistic and cooperative types, discusses implications for work group design and future research, and presents a short mathematical model.

#### ZUSAMMENFASSUNG

Neuere ökonomische Theorien deuten an, dass der Missbrauch von Gruppenakkord-Prämiensystemen durch wechselseitige Überwachung und soziale Sanktionierung gelindert werden könnte. Die vorliegende Arbeit widerspricht dieser Annahme, indem sie zeigt, dass gegenseitige Überwachung und Sanktionierung in Anwesenheit eines Preismechanismus die Motivation zu Kooperation für gewisse Typen von Gruppenmitgliedern hemmen können: Da soziale Belohnungen für Zusammenarbeit auf dem Wunsch nach finanziellem Gewinn aufbauen, kann das Zurückhalten von Anerkennung einem anfänglich kooperativen Teilnehmer weniger bedeuten. Folglich kann gegenseitige Überwachung der Zusammenarbeit in Teams schaden. Der Autor präsentiert Argumente aus der Sozialpsychologie, welche die Unterschiede zwischen individualistischen und kooperativen Typen beleuchten, diskutiert Folgerungen für Arbeitsgruppen-Design und zukünftige Forschung und präsentiert ein einfaches mathematisches Modell.

#### RÉSUMÉ

Les travaux récents en théorie économique avancent la thèse selon laquelle les comportements indépendants au sein d'un groupe soumis à un régime de récompense à la pièce peuvent être réduits par une combinaison de contrôle et de sanctions sociales. Ce travail conteste cette hypothèse en montrant que la présence d'un mécanisme de récompense tant dans le contrôle que dans la sanction contribue à diminuer la motivation à coopérer pour certains types d'agents: cela parce que les récompenses sociales de la coopération qui peuvent apparaître à travers du travail sont potentiellement motivées par un désir de gains pécuniaires; un refus d'accord importe moins pour des agents déjà coopérants. Dès lors, le contrôle mutuel peut diminuer la coopération au sein des équipes. L'auteur présente des données d'un point de vue de la psychologie sociale qui mettent en lumière les différences entre les types de comportements individualistes et coopératifs pour les travaux en groupes et pour les recherches futures ainsi qu'un modèle mathématique.